Problem Statement:

Growing audit requirements have put an administrative burden on organizations across the Department of Defense. As part of these mounting requirements, organizations are required to store and maintain official records for numerous years. Organization and shifting from paper to electronic records and with such a change are storing their records with inconsistent file types and naming convention. Through data science, I wish to explore the ability to recognize documents and underlying metadata to identify the documents, rename them with a consistent naming convention, and save information quickly and accurately. Once identified and processed, unique fields can be quickly searched and audit groups can assess full populations of documentation to help ensure policies and procedures are being followed appropriately.